



Pig oligo annotation/ Version 5

Nov 12th, 2009 / sigenasupport@jouy.inra.fr

Oligo-Set reference :

manufacturer : Aarhus University

name : 25K Pig 70-mer oligonucleotide set

release date : 2003

format : 25K

Organism: pig

Number of oligos: 25 238

Oligo length : 70

Sequences on which oligos have been designed:

44K pig mRNA contigs

"16K Homo sapiens bins"

Antisense strand fixing

47%

Homology filtering

AL \geq 100, %AL \geq 80

"Back-blast" vs hs mRNA

43K seqs

Self-identity filtering

contigID match binID

39K seqs

Selection: most 3'

within bins

15.7K pig mRNA seqs

10K pig mRNA contigs

"8K Mammalian bins"

Antisense strand fixing

41%

Homology filtering

AL \geq 100, %AL \geq 80

"Back-blast" vs mamm. mRNA

8K seqs

Self-identity filtering

contigID match binID

7K seqs

Selection: most 3'

within bins

5.5K pig mRNA seqs

26K pig mRNA contigs

"23K Unknown/pig-specific bins"

Antisense strand fixing

45%

Homology filtering

AL \geq 100, %AL \geq 80

"Back-blast" vs pig ESTs

10K seqs

Selection: max length

within bins

RefSeq filtering

AL \leq 100, %AL \leq 80

5.1K seqs

5.1K pig mRNA seqs

ID GEO:

GPL5972

ID ArrayExpress:

Annotation information

Genome assembly GRCh37, Feb 2009

Gene build Ensembl, Jul 2009

Ensembl version 56, Oct 2009





Specificity analysis

The oligos have been aligned against the transcripts of the species or reference species. The alignment matches are shared between : hit = more than 85% of identity of the oligo length and noise = no hit but more than 20 bp without mismatch : Zhili He's & all 2005

No hit > 85%

6: No noise	8 849
4: 1 noise ≥ 20 & < 30 bp	2 528(*)
3: 1 noise ≥ 30 bp	706(*)
5: Many noise	449

1 hit > 85%

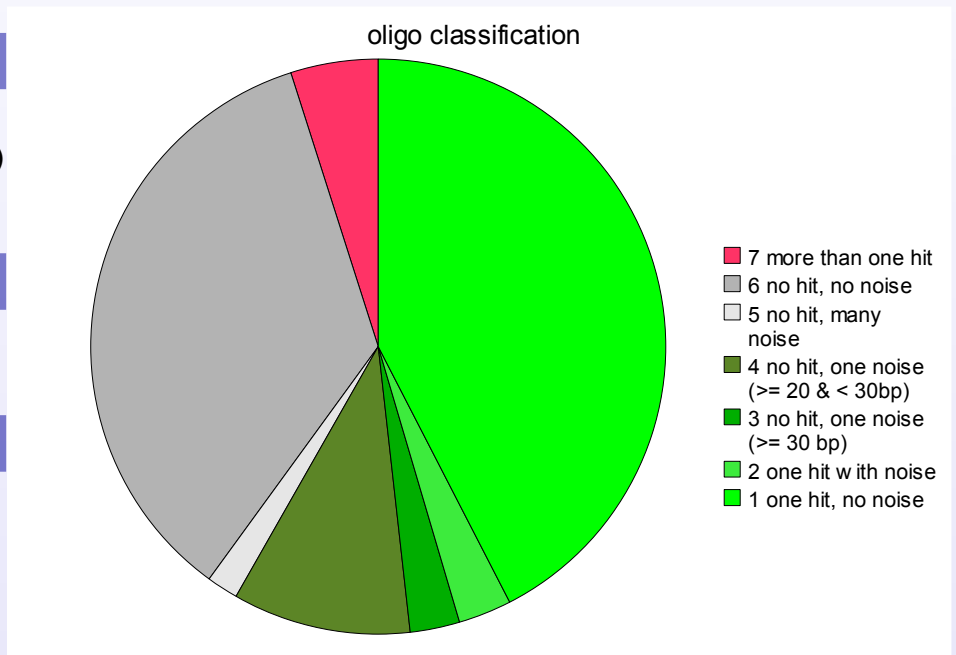
1: No noise	10 713 (*)
2: With noise	752 (*)

More than 1 hit

7: With & without noise	1 241
-------------------------	-------

(*): can be annotated.

Total: **14 699** (green)



Gene coverage

The oligos which can be linked to a gene correspond to **11 279** different genes.

Probes annotated with orthologous Ensembl genes and Gene Ontology

Mouse	13 098	GO Molecular Function	12 682
Rat	12 288	GO Biological Process :	12 111
		GO Cellular Component	12 827

Probes having external reference link

Uniprot/SWISSPROT	14 442
UniGene	12 410
RefSeq_peptide	14 402
HGNC	14 315

